Beatty, I. D., Gerace, W. J., Feldman, A. & Leonard, W. J. (2008). "Technology-Enhanced Formative Assessment: An Innovative Approach To Student-Centered Science Teaching." Proceedings of the Association for Science Teacher Education (ASTE) International Conference, St. Louis, MO.

Technology-Enhanced Formative Assessment:
An Innovative Approach to Student-Centered Science Teaching

Ian D. Beatty, William J. Gerace, Allan Feldman, and William J. Leonard

Scientific Reasoning Research Institute

University of Massachusetts Amherst

Abstract

Technology-Enhanced Formative Assessment (TEFA) is an innovative pedagogical approach to secondary and post-secondary science instruction that uses *classroom response system* technology to teach in accord with educational research findings about effective learning environments. TEFA is built upon four core principles, which we label *question-driven instruction*, *dialogical discourse*, *formative assessment*, and *meta-level communication.* These are implemented in the classroom with an iterative *question cycle.* Mastering TEFA requires developing skill in five different areas: operating the technology, designing effective questions to pose to students, orchestrating whole-class discussion, modeling students and adapting to their needs, and integrating the TEFA approach with curricula and constraints. The details of how teachers learn, assimilate, and adapt TEFA are the object of a current research project.

1. Introduction

In this paper, we describe an innovative pedagogical approach to science instruction at the secondary and post-secondary levels. The approach, named *Technology-Enhanced Formative Assessment* (TEFA), is intended to help students develop deep conceptual understanding, reasoning ability, and problem-solving skills, and to prepare them well for future learning. It identifies four core principles of instruction, which we refer to as *question-driven instruction*, *dialogical discourse*, *formative assessment*, and *meta-level communication*.

TEFA is enacted by structuring class time around an iterative *question cycle* that involves posing a challenging question; giving students time to wrestle with it and discuss it in small groups; collecting answers; holding a class-wide discussion about students' replies, their justifications and merits, and attendant ideas; and closure or wrap-up by the teacher. A *classroom response system* is used to facilitate the cycle and enhance its efficacy.

Section 2 summarizes the literature on classroom response systems. Section 3 identifies the theoretical perspective on science instruction that grounds TEFA, drawing on several different research traditions and orientations. Section 4 defines the TEFA pedagogy by detailing its objectives, its classroom implementation, its core principles, and the role technology plays within it. Section 5 addresses the skills teachers must learn to master TEFA, describes our current research on how teachers learn these skills and assimilate TEFA, and summarizes some preliminary results from that research.

2. Background on Classroom Response Systems

A classroom response system (CRS) is technology that helps teachers poll students during class by collecting their answers to a question that has been posed, instantly aggregating the answers, and presenting a tally of responses (Beatty, Leonard, Gerace & Dufresne, 2004; Banks, 2006; Fies & Marshall, 2006). Modern commercial CRSs are similar in their core functionality (Burnstein & Lederman, 2003): they use simple keypads with buttons as student input devices, which communicate with the teacher's computer via infrared or radio-frequency signals and allow students to indicate their answers to a multiple-choice question. The software instantly tabulates incoming responses and displays a histogram showing the number of students who chose each answer option. Individual student responses remain anonymous to the class, though teachers can usually look them up. Some CRSs provide additional capabilities.

CRS Adoption and Popularity

CRS use has become mainstream in many US universities (Abrahamson, 2006). For example, as of Spring 2004, approximately 8,000 clickers were in use at the University of Massachusetts Amherst, and 6,000 at the University of Colorado Boulder (Duncan, 2005). CRS adoption in K-12 classrooms is more difficult to gauge; manufacturers claim many customers, but the market is so huge that penetration may still be tiny. In early 2005, Abrahamson (2006) conducted a web search and found over 3000 K-12 schools using CRSs.

Where CRSs are used, they are popular. Most university students are enthusiastic about using "clickers" (Roschelle, Abrahamson & Penuel, 2004a; Roschelle, Penuel & Abrahamson, 2004b; Fies & Marshall, 2006). Fagen, Crouch, and Mazur (2002) identified 384 teachers who had tried CRS-based instruction with an approach similar to *Peer Instruction* (see below), of whom 303 "definitely" intended to use it again. Data on CRS popularity at the K-12 level is lacking, but in our work with secondary school teachers, we have found that most students and many teachers are enthusiastic about CRS use.

CRS Pedagogies

Few authors of CRS-related papers explicitly define their pedagogy, and "teaching with a response system" is frequently treated as if it *were* a form of pedagogy. However, a CRS is merely a tool, and like any tool it may be put to many different purposes and used well, ineffectively, or even counterproductively. Thus, it is crucial that CRS advocates and researchers clearly define the pedagogic practices that they associate with "CRS use".

We now summarize the three published, explicit, coherent pedagogies for CRS use we are aware of. All three have their roots in university physics instruction; Fies and Marshall (2006) have observed that CRS use and research into it has disproportionately occurred within the discipline of physics.

Peer Instruction and ConcepTests

The best-known pedagogy for teaching with a CRS is Eric Mazur's *Peer Instruction* (Mazur, 1997). Mazur suggests regularly inserting CRS-administered *ConcepTests* — multiple-choice conceptual questions about the material being taught — into the lesson. If a significant number of students answer incorrectly, the class is asked to discuss the question among themselves and then answer again. Mazur argues that this methodology increases student engagement, improves learning, provides the teacher with feedback about student understanding, and promotes knowledge "diffusion" between students. Quantitative evidence, primarily from pre/post testing with the Physics *Force Concept Inventory* (Hestenes, Wells & Swackhammer, 1992), supports his assertion that Peer Instruction improves student understanding.

Assessing-to-Learn

Simultaneously, the University of Massachusetts Physics Education Research Group (UMPERG) was developing its own CRS pedagogy. At first unnamed (Dufresne, Gerace, Leonard, Mestre & Wenk, 1996; Wenk, Dufresne, Gerace, Leonard & Mestre, 1997), they later referred to it as *Assessing-to-Learn* (A2L; Dufrense, Gerace, Mestre & Leonard, 2000) or *Question-Driven Instruction* (Beatty, Leonard, Gerace & Dufresne, 2006b). Dufresne et al. (1996) define four "broad educational objectives" of the approach:

> 1) Students should know and understand definitions, terminology, facts, concepts, principles, operations, and procedures; 2) Students should be able to communicate what they know to others; 3) Students should know how to apply what they have learned to analyze situations and solve problems, extending this ability to increasingly complex situations; and 4) Students should develop the ability to evaluate critically the usefulness of various problem-solving approaches… We do not take for granted that students will acquire or enhance these habits of mind working independently outside of class. (p. 12)

Dufresne et al. (2000) later elaborated on these goals, identifying twelve specific "beneficial habits of mind" to help students develop and five "stages of cognitive development" to address in order to help students construct well structured, robust, transferable knowledge structures and problem-solving strategies. They also connected A2L with formative assessment, saying that "it informs teachers about what students think; it informs students what their classmates think; it informs individuals what they themselves think" (p. 11).

The primary mechanism by which A2L addresses these goals is the "question cycle", which we will describe in Section 4. An important difference between *Peer Instruction* and A2L is that Mazur's ConcepTests are intended to be inserted intermittently within more "traditional" instruction in order to enhance and guide that instruction, whereas the A2L question cycle forms the basic structure of class activity, with "micro-lectures" or other direct instruction inserted when needed and motivated by the questions and discussion.

OSU Question Sets

More recently, the Physics Education Research Group at The Ohio State University (OSU) has been developing a methodology for CRS use that focuses on *sets* of related questions working together to develop understanding of one concept (Ohio State University Physics Education Research Group, 2007). While the general idea of using coordinated question sets as a coherent instructional "unit" is not new (Dufresne et al., 1996; Beatty, Gerace, Leonard & Dufresne, 2006), Reay, Bao, Warnakulasooriya, and Baugh (2006) proposed and extensively tested two specific design patterns for such sets. One pattern, called "easy-hard-hard", is a series of three questions about the same concept. The first question is an easy "warm-up" designed to build confidence; it typically has little discussion. The second question is difficult: it pushes the limits of students' understanding of the concept, and is intended to elicit a broad spectrum of answers and lead to extensive discussion. The third question is also difficult, in the same way as the second, but with different surface features. It reveals to students and the teacher whether students have learned what they were intended to from the second question by whether they can apply it in a different context.

The OSU group has developed a second question-set methodology that they call “rapid-fire”, in which a series of moderately difficult questions present one concept in a variety of contexts (Ohio State University Physics Education Research Group, 2007). They are still studying this approach, but believe it might be more effective than “easy-hard-hard” sets for lower-achieving students.

Empirical Research on CRSs

We are aware of two reviews of research on the use of CRSs. The first, published in 2004 (b) by Roschelle, Abrahamson, and Penuel, examined 26 CRS studies. They found 16 that report increased student engagement; 11 that report increased student understanding of subject matter; 7 that report increased student enjoyment of class; 6 that report improved group interaction; 5 that claim using a CRS helped students gauge their own understanding; and 4 that report an improvement to teachers’ awareness of student difficulties. While they concluded that the studies are indicative of “a real and important phenomenon at hand” (p. 3), they said that

> none of the available studies rises to the present specification of “scientifically based research” that would allow inferences about causal relationships or that could form the basis for estimating the magnitude of the effect. (p. 3)

They also said

> our review found that existing research does not connect with the larger research base in education or psychology, which could be used to create an explanatory theory or model. (p. 3)

The second review, by Fies and Marshall (2006), reached largely the same conclusions. They added that “Amongst the most commonly stated benefits of CRS use are improved attendance and participation, which may to some degree be attributable to the practice of making part of the course grade dependent upon CRS input” (p. 105).

A strong conclusion of the reviews is that students like using CRSs. This may be their most important finding, given Koballa and Glynn’s assertion that “Science learning experiences that are fun and personally fulfilling are likely to foster positive attitudes and heightened motivation toward science learning and lead to improved achievement" (2007, p. 94). In addition, CRS use may increase because teachers like using the technology (Fagen, Crouch & Mazur, 2002), and they are the ones who usually select teaching tools.

Theorizing CRSs

The only significant effort we are aware of to provide an overarching theoretical framework for CRS-based teaching is that of Roschelle et al. (2004a; 2004b). Based on an analysis of reported CRS benefits, they suggest focusing on four main constructs to connect CRS-based instruction to the broader educational research literature:

- formative assessment,
- driving discussion by important conceptual contrasts,
- shifting to mastery-oriented motivational incentives, and
- harnessing diversity for generativity.

They do not believe that CRS-based instruction *necessarily* does these things, but rather that these constructs are useful to understand what *can* happen.

3. Theoretical Framework

Our theoretical framework for TEFA builds on research in four areas: cognition and conceptual learning, social aspects of learning, student attitudes and motivation, and formative assessment.

The Cognitive Dimension

To teach science largely means to develop and refine students' understanding of the concepts of science. Therefore, *constructivism* and the *conceptual change research tradition* are central to our perspective. We use the term *constructivism* to mean that

- knowledge is constructed, not transmitted;
- the construction of knowledge requires purposeful and effortful activity;
- prior knowledge impacts the learning process; and
- initial understanding is local, not global. (Gerace, 1992)

These premises align with the fundamental insights into conceptual learning of science shared by the majority of cognitive perspectives (Scott, Asoki & Leach, 2007).

Following Posner, Strike, Hewson, and Gertzog (1982), we see learning science as a complex process of growth and reorganization of an individual student's "conceptual ecology". The development of scientific understanding requires integrated change to an interlocking set of ideas, and pre-existing beliefs must be explicitly addressed. Because experts and novices differ in the organization, not just the extent, of their knowledge (Larkin, 1979; Chi & Glaser, 1981; Glaser, 1992), effective science instruction must help students reflect upon and *structure* their knowledge appropriately.

Recent research suggests that when students seem to exhibit a gap in their knowledge, they frequently *possess* the requisite knowledge elements but fail to *access* them in the specific context at hand (Redish, 2003; Hammer, Elby, Scherr & Redish, 2004; Dufresne, Thaden-Koch, Gerace & Leonard, 2005; Scherr, 2007). We are therefore concerned with the range and nature of the contexts in which students explore new science knowledge, and value learning experiences in which students must search their accumulated knowledge, weighing alternatives and choosing elements to apply to specific situations.

What knowledge students even attempt to access while answering questions and solving problems, or seek to develop while learning, is strongly constrained by their *epistemological framing* of a situation (Elby, 2001; Hammer & Elby, 2003). A person's *frame* of the moment is their answer to the questions "What is it that's going on here, and what should I be trying to do?" Consequently, we believe that attention to how students frame their participation in learning is vital to effective instruction.

The Social Dimension

Within the conceptual change tradition, the primary role of classroom discourse is to direct students' thinking and provide material for them to think about. The *sociocultural research tradition* identifies other crucial roles for language in science instruction. Carlsen (2007), extending Sutton's (1998) work, articulates three distinct ways of conceptualizing language in science and science teaching: as a system for transmitting information, as an interpretive system for making sense of experience, and as a tool for participation in communities of practice.

According to Bakhtin (summarized in Wertsch, 1991, pp. 93-118), learning science or mathematics involves developing fluency in the *social language* of the discipline: the language, concepts, norms, and genres for communication used by the discipline's practitioners. The assemblage of social languages that a person knows comprises a "toolkit" of ways of knowing and thinking. Lemke (1990) goes farther, asserting that "learning science" means "learning to *talk* science". He sees the content of science curricula as *thematic patterns*, networks of semantic relationships between words and their corresponding meanings, which must be learned through discourse experiences.

The nature of scientific social languages is qualitatively different from the nature of everyday social languages, in both ontological and epistemological ways (Mortimer & Scott, 2003; Carlsen, 2007). Science social languages can view the same phenomena in very different ways from everyday ones, leading to what researchers in the cognitive change tradition would call "misconceptions" or "preconceptions" as well as to a disconnect between real life and "what we learn in the classroom". To learn science, students need help recognizing and resolving conflicts between alternative social languages.

This is also important because, according to Vygotsky (1987), ideas are first encountered and "rehearsed" through communication on the *social plane*. As an individual reflects upon and makes sense of these, the social tools for communication become internalized and provide the means for individual thinking. Therefore,

> any sequence of science lessons, which has as its learning goal the meaningful understanding of scientific conceptual knowledge, must entail both authoritative and dialogic passages of interaction. (Mortimer & Scott, 2003, p. 2)

However, according to Scott and Mortimer (2006), "dialogic interactions are notably absent from science classrooms around the world" (p. 2). Accordingly, we take the practice of genuinely dialogic discourse to be vital for effective science instruction.

The Attitudinal Dimension

Understanding what motivates students to learn is also important, so we turn to *research on student attitudes and motivations*. Koballa and Glynn (2007) identify four general theoretical orientations within the research literature on student motivation: *behavioral*, which focuses on incentive and reinforcement; *humanistic*, which focuses on students' personal growth and desire to self-actualize; *cognitive*, which focuses on students' goals, plans, expectations, and attributions; and *social*, which focuses on students' identities and interpersonal relationships. The cognitive, humanistic, and social orientations all inform our perspective. The cognitive orientation is dominant, in that it grounds our thinking about the *mechanisms* of motivation and behavior and the avenues we have for affecting them. The humanistic orientation informs our goals, and the social orientation helps us understand and make use of the ways social dynamics and context impact individual attitudes.

According to Koballa and Glynn (2007), a key construct for understanding student motivation is *intrinsic motivation*. "Motivation to perform an activity for its own sake is intrinsic, whereas motivation to perform it as a means to an end is extrinsic" (p. 89). Their literature synthesis found that five factors influenced students' degree of intrinsic motivation: teacher expectations, goal-directed behavior, self-determination, self-regulation, and self-efficacy. Students' feelings of *self-determination* — belief that they have some control over their behavior, and can make choices regarding it — tends to result in higher levels of achievement, as well as better emotional adjustment. Sustained lack of self-determination can result in *learned helplessness*, a condition in which students are reluctant to engage in learning activities because they have no expectation of success. *Self-regulation* of learning means continually monitoring one's own learning progress relative to chosen goals and adjusting learning activity accordingly. It consists of *regulatory strategy use* for planning activity and monitoring progress, and also of *cognitive strategy use* for organizing and elaborating the material being learned. Students' willingness to self-regulate is weakened by a poor sense of control and self-determination.

The *theory of reasoned action* (Koballa & Glynn, 2007) posits that beliefs determine attitudes and attitudes shape behavior. Most research about student attitudes in science education focuses on attitudes towards science; we are more urgently interested in attitudes towards learning, instruction, and classroom behavior, but we expect the underlying psychological dynamics to be the same.

Formative Assessment

Formative assessment is any assessment that "contribute[s] to student learning through the provision of information about performance" (Yorke, 2003, p. 478). In practice, a teacher sets up a situation that elicits information about a student's understanding or knowledge, and then the teacher and/or student interprets and acts upon that information (Sadler, 1989; Black & Wiliam, 1998a; Bransford, Brown & Cocking, 1999; Bell & Cowie, 2001; Black & Wiliam, 2005).

According to Black and Wiliam (1998b), "innovations which include strengthening the practice of formative assessment produce significant, and often substantial, learning gains" across ages, school subjects, and countries — gains "larger than most of those found for educational interventions" (p. 140). Formative assessment is particularly beneficial for traditionally “low achieving” students, with potential to help narrow the achievement gap between students from different socioeconomic strata (Black & Wiliam, 1998b; Stiggins, 2002). Formative assessment can elicit richer classroom discourse and help students become more engaged and motivated (Gallagher, 2000), can help students become aware of the limits of their understanding and the actions they can take to progress (Ramaprasad, 1983; Sadler, 1989), and can catalyze significant teacher learning (Bransford, Brown & Cocking, 1999; Black et al., 2002).

Qualities of Effective Learning Environments: A Synthesis

How People Learn (Bransford, Brown & Cocking, 1999) synthesizes a broad survey of educational research to argue that effective learning environments should be *student-centered*, *knowledge-centered*, *assessment-centered*, and *community-centered*. A *student-centered* learning environment treats students as individuals, coaching them from their varied initial states to the intended learning goal by whatever unique trajectory each requires, taking into account their initial knowledge and perceptions, their culture, their language use, and their ongoing and very personal process of sense-making. Student-centered teachers also engage in formative assessment practices by

> attempting to discover what students think in relation to the problems on hand, discussing their misconceptions sensitively, and giving them situations to go on thinking about which will enable them to readjust their ideas. (pp. 133-4)

They also respect the language practices of their students as a basis for further learning (p. 135).

A *knowledge-centered* learning environment treats knowledge as a rich, interconnected structure that must be organized and refined as it is expanded.

> [K]nowledge-centered environments also include an emphasis on sense-making — on helping students become metacognitive by expecting new information to make sense and asking for clarification when it doesn’t. (p. 137)

An *assessment-centered* learning environment weaves *formative assessment* deeply into the fabric of instruction, providing continual, detailed feedback to guide students’ learning and teachers’ teaching. In a nod to CRSs (as well as other instructional technologies), the authors note that “Teachers have limited time to assess students’ performances and provide feedback, but new advances in technology can help solve this problem…” (p. 142)

A *community-centered* learning environment recognizes that students belong to communities of co-learners at the course, program, institution, and society levels, and promotes constructive interaction between individuals to further learning. It recognizes that

> learning seems to be enhanced by social norms that value the search for understanding and allow students (and teachers) the freedom to make mistakes in order to learn. (p. 145)

Taken together, these four qualities can help us evaluate instructional environments for their alignment with educational research. They do not, however, prescribe what teachers must *do* to produce these qualities. For that, we need a pedagogy.

4. Technology Enhanced Formative Assessment

Technology-Enhanced Formative Assessment (TEFA) is our pedagogical approach for teaching science with a classroom response system. It originated in our early efforts to extend and adapt general research-based physics teaching strategies (e.g., Gerace, 1992) to capitalize on the affordances of classroom response technology, and grew through our experiences teaching with it, researching it, and mentoring others in its use.

Instructional Goals of TEFA

We developed TEFA for two general purposes: to help students develop expertise in science content, and to help prepare students for future learning (Bransford & Schwartz, 1999; Schwartz & Martin, 2004). In alignment with conceptual change literature, TEFA aims to help students grow contextually robust, transferable conceptual frameworks that are well reconciled with their experiences, perceptions, and prior understandings, and develop concept-based problem solving and model-based reasoning skills. In alignment with sociocultural learning literature, TEFA aims to engage students in rich, dialogical discourse about scientific ideas and their applications. In alignment with student motivation literature, TEFA aims to explicitly confront students' beliefs and attitudes, communicate high teacher expectations, and scaffold self-directed, self-regulated learning habits. Students must be taught to recognize and seek well-structured knowledge, to participate in productive modes of discourse, and to attentively self-regulate their learning.

The Question Cycle

The most tangible aspect of TEFA practice is the *question cycle*. In the classroom, TEFA is implemented by structuring whole-class interaction around an iterative, CRS-supported cycle of question-posing, answering, and discussing (Dufresne et al., 1996). The essential phases of the cycle are:

1. Pose a challenging question or problem to the students. (In TEFA, we do not teach and then ask questions about what was taught; we ask questions first, and use them as a context for sense-making and direct instruction.)
2. Have students wrestle with the question — alone, in small groups, or both in succession — and choose a response.
3. Use a CRS to collect responses, even from students who are unsure, and display a histogram of the aggregated responses.
4. Elicit from students as many different reasons and justifications for the chosen responses as possible, without revealing which (if any) is or are correct. In the process, draw out students' reasoning and vocabulary, expose students to each other's ideas, and make implicit assumptions explicit.
5. Develop a student-dominated discussion on the assumptions, perceptions, ideas, and arguments involved. Help students formulate their ideas and practice "talking science", find out why they think what they do, and gently increase their understanding. (In practice, phases 4 and 5 blend together.)
6. Provide a summary, direct instruction, meta-level comments, segue to another question, or whatever other closure seems warranted, guided by the detailed information just revealed about students' thinking. The class is now well primed to receive the message, appreciate its relevance, and integrate it with other knowledge.

Questions can build upon each other or function together as sets in order to develop students' understanding. Demonstrations, a second answer-collecting round after some discussion, and other elaborations may be included as appropriate. We find that iterating through this cycle three or four times in 50 to 60 minutes of TEFA instruction is appropriate; a higher rate does not give students enough time to really engage, ponder, discuss and listen, and practice speaking to the degree that we intend.

The TEFA question cycle is flexible and rich enough to be a regular, perhaps dominant, part of science instruction. It does not, however, address every instructional need. Most teachers will need to include complementary course components such as pre-class reading (for initial exposure to ideas), post-class homework (for more intensive problem-solving work), group projects (for extended explorations of ideas), and laboratory exercises (for hands-on learning opportunities).

Uses for TEFA Questions

TEFA questions and the question cycle can serve multiple ends. A teacher can achieve any one or more of the following general objectives with a question:

- learn about students' knowledge, thinking, and perceptions;
- help students become more aware of their own knowledge, thinking, and perceptions;
- help students become cognizant of other students' knowledge, thinking, and perceptions;
- set up subsequent instruction;
- catalyze small-group discussion and peer learning;
- provoke, open, motivate, ground, and contextualize whole-class discussion of a topic; and
- precipitate student insights and realizations.

The following list provides examples of some more specific ways that TEFA questions and the question cycle can play within an instructional plan.

- Status check: During instruction, poll students for their degree of confidence in their understanding of a specific topic.
- Exit poll: At the close of a class session, poll students to find out which of several concepts covered that day they most want to spend more time on.
- Assess prior knowledge: Elicit what students already know, think, believe, or perceive about a topic or idea before formally addressing it in class.
- Provoke thinking: Ask an intriguing and challenging (but approachable), question to "open up" a new topic or subject, get students engaged and thinking about it, and provide context and shared experience for subsequent learning.
- Elicit a misconception: Lead students to manifest a specific common misconception or belief that may hinder their learning, so that it may be articulated, examined, and dispatched.
- Exercise a cognitive skill: Drive students to engage in a specific type of cognitive activity or "habit of mind" (e.g., seeking alternative representations, comparing or contrasting two situations, categorizing and classifying cases, or strategizing and planning a solution; Dufrense et al., 2000) in order to strengthen that habit.

- Build conceptual structure: Hone, link, or extend a concept by challenging students to identify its limits of applicability, differentiate it from a similar concept, recognize a relationship with a distinct concept, or apply it in a new context.
- Stimulate discussion: Provoke dialogical whole-class discussion with a "disputable" question having multiple reasonable or defensible, but not obviously right, answers.
- Induce cognitive conflict: Deliberately bring students to the realization that two of their beliefs, perceptions, ideas, interpretations, or models are in conflict with each other, thus creating a "teachable moment".
- Predict a demonstration: Ask students to predict the outcome of a demonstration or experiment, and commit to that prediction, so that they will be more attentive and learn more when their prediction is either confirmed or disconfirmed.
- Test capability: Determine whether students have developed the capacity to solve a particular kind of question.
- Demonstrate success: Build students' confidence and help them to recognize their own progress by posing a question that most should be able to answer successfully.

This list should serve to demonstrate TEFA's flexibility and help teachers avoid getting "stuck in a rut" by repeatedly using the same types of questions in the same overall pattern.

The Four Principles of TEFA

Mechanically following the question cycle does not, by itself, constitute "doing TEFA". Just as a CRS is a tool for conducting the question cycle, the question cycle is a vehicle for realizing TEFA's four key principles:

1. Motivate and focus student learning with *question-driven instruction*.
2. Develop students' understanding and scientific fluency with *dialogical discourse*.
3. Optimize teaching and students' learning with *formative assessment*.
4. Help students cooperate in the learning process and develop metacognitive skills with *meta-level communication*.

Question-Driven Instruction (QDI)

The first TEFA principle is "Motivate and focus student learning with *question-driven instruction*" (QDI). It follows from an appreciation of constructivism and the body of research on conceptual change and knowledge access. It is realized in the question cycle by the placement of question posing at the beginning of the cycle, framing all that follows, with the "closure" phase

— where most "direct teaching" will occur — at the end. According to Bransford, Brown, and Cocking (1999): "Ideas are best introduced when students see a need or a reason for their use — this helps them see relevant uses of knowledge to make sense of what they are learning" (p. 139).

TEFA positions learning within the context of students' encounter with challenging, conceptually rich, preferably meaningful questions that provide context, motivation, and direction to students' sense-making efforts. Questions are used to set up fertile learning situations and to instigate learning, not just to assess or consolidate the results of previous instruction.

Dialogical Discourse (DD)

The second TEFA principle is "Develop students' understanding and scientific fluency with *dialogical discourse*" (DD). It follows from the body of research on sociocultural learning. It is realized in the question cycle in both the small-group and whole-class discussion phases. Whole-class discussion is generally begun by identifying different answers that have been selected and eliciting arguments or explanations for them, *before* scrutinizing the validity of any response, as a strategy for increasing the dialogicity of the discourse.

Discussion within TEFA is intended to have several effects:

- to clarify thought through the process of articulation and externalization;
- to expose students to different points of view and lines of thinking;
- to promote analysis and resolution of disagreements;
- to supply stimuli, context, and tools for individual sense-making; and
- to provide practice speaking the social language of science.

We believe that the majority of student learning in TEFA happens during the whole-class discussion phase of the question cycle. Thus, orchestrating high-quality discourse — interactive, dialogical, and thematically rich — is a top priority for teaching with TEFA.

Formative Assessment (FA)

The third TEFA principle is "Optimize teaching and students' learning with *formative assessment*" (FA). It follows directly from the research literature on the effectiveness of formative assessment.

FA is realized in the question cycle in multiple ways. Students learn about what they do and don't understand by their ability to answer the posed questions and by how their responses compare to their peers'. They learn more about the extent of their own understanding in the process of trying to articulate it coherently and convince others, and yet more when the teacher provides well tuned, apropos, prescriptive feedback as part of the closure phase. Teachers learn about their students' understanding, perceptions, assumptions, and reasoning from the histogram of question responses, and in more detail from what students say as they defend, explore, and contrast their ideas.

Meta-Level Communication (MLC)

The fourth TEFA principle is "Help students cooperate in the learning process and develop metacognitive skills with *meta-level communication*" (MLC). It is motivated by literature on student attitudes and motivation, though there is little research on MLC itself (e.g., Wilson, 2006). We identify three categories of MLC that are significant to TEFA: meta-narrative, metacognitive talk, and metacommunication. (*Metadiscourse*, which simply means "talk about talk" [Lemke, 1990], is a more general concept from semantic analysis that overlaps with these three.) We can say that most of the talk in a class is discourse about the subject, but MLC is discourse about *learning* the subject.

Meta-narrative is communication about the purpose, design, and unfolding of the course from a "higher" perspective. Its purpose is to make students more consciously aware of what is going on in the class and why it is happening, so that they may frame their activity appropriately, focus on the most salient aspects, and actively seek the right kind of realizations.

Metacognitive talk is communication about thinking, learning, knowledge, and similar cognitive or epistemological issues. Its purpose is to improve students' understanding of these things so that they become more aware learners and can make wiser choices about their learning actions.

Metacommunication is communication about communication. Its purpose is to refine the communication in the classroom and help students participate more consciously and efficiently in it. Seen through the lens of constructivism, communication is inherently error-prone: the recipient of a message attaches his or her own meanings to the words and constructs an interpretation of the message that may or may not align well with the sender's intention. This difficulty is exacerbated in teaching, where the student is unfamiliar with much of the language being used. It follows that awareness of the potential for miscommunication and proactive monitoring of message fidelity — such as asking or saying something in multiple ways, or actively considering alternative interpretations of a message — improve communication efficiency. The more active students become about aggressively improving communication, the more effective instruction can be (Gerace, 1992ccr).

These three categories of MLC overlap. A statement about the teacher's purpose in saying or asking something may be both meta-narrative and metacommunication. Similarly, explanation of the purpose of some part of the course in terms of its role in the learning process could be meta-narrative and metacognitive talk. Often, we find that different kinds of MLC chain together, with a meta-narrative or metacommunicative comment leading to a more extended metacognitive discussion of some issue.

Within TEFA, MLC serves as the primary tool for affecting students' beliefs, attitudes, and motivation, and thus for altering their behavior patterns. When we practice TEFA, we do not *train* students to engage in the question cycle the way we wish. Instead, we *invite* them, attempting to make them consciously aware of their choices for learning activity and the ramifications those choices can have, thus enhancing self-determination and inviting self-regulatory learning. We suggest and encourage self-regulatory strategies, thus scaffolding the development of self-regulation. We indicate ways to improve success at communication and learning, thus enhancing self-efficacy. We challenge beliefs about teaching, learning, and science by suggesting and defending alternative interpretations, thus influencing attitudes and behavior. And we help students select appropriate epistemological frames for their participation by meta-communicating the nature and purpose of activities.

TEFA employs MLC to improve *learning* by increasing the efficiency of the instructional process, and to improve *the learner* by promoting and scaffolding student development of more productive learning beliefs, attitudes, and behaviors. These mirror TEFA's twin goals of helping students develop expertise in the science subject being taught, and helping to prepare them for future learning.

Synergy Between the Principles

TEFA's four principles are not a collection of independent pedagogical exhortations. They interlock and reinforce each other. In the absence of any one, the other principles become less effective and more difficult, and TEFA unravels. For example, successful question-driven instruction requires tuning questions to students' "zone of proximal development" (Vygotsky, 1978) and scaffolding their efforts just enough to help them succeed, without bypassing the confusion, struggle, and conflict; formative assessment helps a teacher gather information to tune correctly. Formative assessment requires gathering data about what students are thinking, and why; question-driven instruction and dialogical discourse provide complementary data sources on this. Dialogical discourse requires a context and focus, and works better when students have at least provisionally committed to some position; question-driven instruction supplies those. Question-driven instruction, dialogical discourse, and formative assessment are all aided by students' active, well-intentioned, well-informed cooperation; meta-level communication helps cultivate that. The four principles enhance each other synergistically across phases of the question cycle.

The Role of Technology in TEFA

Nothing about TEFA *in principle* requires a CRS. However, using the technology to assist with TEFA offers several benefits that enhance, and in some contexts make realistic, what TEFA prescribes.

A crucial feature of CRSs is that they provide both *anonymity* and *accountability* (Roschelle, Abrahamson & Penuel, 2004a; Fies & Marshall, 2006): students can be held accountable for answering questions, but the actual answer each student has chosen is not revealed to other students and is not immediately obvious to the teacher. True anonymity is difficult to arrange by raised hands (even with heads down), color-coded cards, or other means. Equivalent anonymity can be achieved via paper forms, but this is too slow for real-time formative assessment.

Additionally, a CRS allows collecting answers from *all* students in a class, rather than just a few who speak up or are called upon. This means all students can benefit from the cognitive act of choosing and committing to an answer. It also means the teacher gains better data about students' thinking. "Increased student engagement and participation" is one of the two most commonly reported findings of CRS implementation studies (Roschelle, Abrahamson & Penuel, 2004a; Fies & Marshall, 2006).

We conjecture that CRS technology provides an additional benefit in that the act of pushing a button and definitively submitting an answer, with no waffling or qualification possible (until the discussion phase, of course), amplifies the psychological benefit of forcing students to "pick a side". Once students have committed to an answer, whether or not they are confident in it, we believe they attend to subsequent discussion and resolution of the matter in a different and more attentive way. The answer they have selected is now "their" answer, for good or ill, and they want to see how it fares.

The CRS histogram showing the distribution of students' answers also adds value to the process. It is not just a way to find out how many picked which answer; it is also, as Roschelle, Abrahamson, and Penuel (2004a) note, a "high contrast display that drive[s] productive discourse" (p. 28). It makes differences in students' positions starkly obvious; one glance strongly conveys whether the class is in agreement (a single peak), generally undecided (a uniform or random spread), or highly polarized (two distinct peaks). It also serves as a point of focus and reference for subsequent discussion and instruction. A histogram communicates the same information as a list of numbers could, but in a way that is more forceful and easier to digest — another advantage of CRSs over hand-raising and flash card methods.

Finally, a CRS can record the data of students' individual and collective responses for subsequent analysis. This can help a teacher diagnose class-wide or individual student needs, or self-evaluate his or her own instruction.

5. Teacher Learning of TEFA

TEFA is a complicated and comprehensive pedagogical approach. For most teachers, adopting it means assuming new roles, developing new skills, focusing on new outcomes, adopting new class planning procedures, establishing new classroom norms and patterns, and fostering new student attitudes. This is not easy.

Our current project, *Teacher Learning of Technology-Enhanced Formative Assessment* (TLT, NSF TPC-0456124), is devoted to building a rich model of teacher learning of TEFA: what difficulties teachers have, what obstacles impede them, what insights and interventions help them progress, what stages they advance through, what distinct kinds of trajectories they follow, and so on. The study has three main goals:

1. To better understand teacher learning of the TEFA pedagogical approach;
2. To better understand effective and efficient methods of teacher professional development in TEFA; and
3. To develop tools and techniques for the evaluation of teachers' TEFA mastery and implementation fidelity, of suitable design and quality for use in a controlled, randomized study of the effects of TEFA on student learning.

This section will present a synopsis of past conclusions about how teachers learn TEFA, outline the design of the current project, and summarize some preliminary findings.

Previous Conclusions About Teacher Learning of TEFA

Mastering TEFA is a slow process. Most teachers require about three years of concerted use to become fully comfortable with TEFA (Feldman & Capobianco, in press), so that it is part of their "way of being a teacher" (Blum, 1999; Davis, Feldman, Irwin, Pedevillano, Capobianco & Weiss, 2003). It is possible, but unproven, that better professional development techniques and curriculum could reduce that time somewhat.

One of the reasons why mastering TEFA is a slow process is that teachers must develop skills in five distinct areas in order to "master" the pedagogy (extending Feldman & Capobianco, in press). They are:

1. operating the CRS technology and managing the logistics of the classroom;
2. designing (or locating and adapting) quality questions;
3. eliciting and orchestrating productive dialogical discourse in the classroom;
4. interpreting, modeling, and dynamically adapting to students in real-time; and
5. integrating TEFA practice with non-TEFA instructional activities, curriculum materials, frameworks and standards, standardized exams, and other constraints and contextual factors.

Operating the Technology

The first and most obvious skill area that new TEFA practitioners must address is learning to use the CRS technology itself. This includes connecting the hardware, operating the software, teaching students how to use their transmitters, and troubleshooting problems. It also includes resolving classroom logistical issues such as organizing and placing student transmitters for orderly pick-up and return, situating a video display or data projector and screen, arranging furniture to support both small-group and whole-class discussion, and navigating a possibly cluttered classroom to operate the CRS software while interacting naturally with students.

Most teachers surmount these challenges within a few weeks. They do not need to understand all the details and extra features of the CRS system, and most do not try. Rather, they develop a simple routine that serves their immediate teaching needs, learning only the technology features it requires. Thereafter, they explore additional features and develop additional skills at their convenience, motivated by curiosity or dissatisfaction and a growing sense of competence (Feldman & Capobianco, in press).

Designing Questions

As teachers become comfortable operating the essential functions of a CRS, their attention turns to the second area skill area, question design. They quickly realize that constructing questions to precipitate student learning, provoke dialogical discourse, and support formative assessment is more difficult than they had anticipated. It requires pedagogical content knowledge (van Driel, Verloop & de Vos, 1998), knowledge of students, careful crafting, and anticipation of contingencies. An arsenal of strategies and tactics for question creation exists (Beatty, Gerace, Leonard & Dufresne, 2006a), but these take time to understand and assimilate, and only partially reduce the difficulty of the task.

For most teachers, learning to develop satisfactory questions is the primary focus of attention for the majority of their first year of TEFA. The work is considerably easier in subsequent repetitions of the same course, since the teacher now has a library of field-tested questions to work from and knowledge of how they will play out in class.

Orchestrating Discourse

The third skill area also confronts teachers early in their learning of TEFA. Shortly after grasping the complexity of question creation, most teachers become dissatisfied with the classroom discussions that result, but are not sure what to do about it. For many teachers, stimulating and orchestrating productive dialogical discourse seems to be more deeply challenging than question creation.

We see three likely reasons for this. The first is that making good discourse happen requires cooperation from the students, and thus is not under the direct control of the teacher. The second reason is that orchestrating discussion must be done “in the heat of battle” under heavy cognitive load, with little time to consider tactics that are not yet “second nature”. The third reason is that good classroom discussion depends strongly on several things: the nature of the question used, the mix of ideas that students bring to it, students’ attitudes and expectations, and how the teacher starts and steers the discussion. Teachers often report an occasional and very rewarding experience of strong student discussion among many disappointing or semi-satisfactory experiences, and only gradually discern the factors that made those few go so well. (Some teachers have a very different experience, at least with certain classes: trying to keep the students from talking so much that nothing else gets done.)

Modeling and Adapting to Students

The fourth TEFA skill area is interpreting, modeling, and dynamically adapting to students in real-time — what we colloquially call "getting inside students' heads, and then knowing what to do about it." Teachers rarely identify this explicitly as a skill area they want to develop, at least during the first year or so of TEFA practice. We believe that all teachers wrestle with it to some degree, but intuitively and under the guise of learning what kinds of questions and discussion tactics "work". Some research has found that using a CRS for formative assessment can improve a teacher's mental models of students, at least by increasing their awareness of common student difficulties (Roschelle, Abrahamson & Penuel, 2004a).

We identify this as a distinct TEFA skill area for two reasons. First, as teachers practice formative assessment and succeed at eliciting information about students they had not previously known, they inevitably confront the question of "So now that I know that, what do I do about it?" They find that making sound real-time teaching adjustments — which we call *agile teaching* (Beatty et al., 2006b) — is hard. Wiliam (2007) says

> This study [just described] seems to indicate that collecting data if one cannot do anything with it is counterproductive… [E]ven when teachers do manage to use information about student achievement to adjust or individualize their instruction, teachers may lack the ability to do so effectively. (p. 10)

The second reason we identify this as a distinct TEFA skill area is that the few highly advanced TEFA practitioners we've known — for whom operating the technology, designing questions, and orchestrating discourse have become second nature — shift their focus squarely onto their students, trying to build ever more accurate, comprehensive, and sophisticated mental models of their students' knowledge and skills. Roughly speaking, TEFA learners' attention seems to move from the technology (the tool) to the questions (the plan), to the interactions in the classroom (the execution), to the machinations inside students' heads (the target).

Integrating with Other Curriculum and Constraints

The fifth TEFA skill area is integrating TEFA practice with non-TEFA instructional activities, curriculum materials, frameworks, standardized exams, and other constraints and contextual factors. TEFA practice does not exist in a vacuum; constraints must be lived within and other demands must be accommodated. The skills required to do so and the acquisition of these skills are highly idiosyncratic, depending sensitively on the details of a teacher's outlook and context. We have seen two science teachers from the same school, or two teachers of exactly the same subject, grade, and level at two similar schools, have very different perceptions of the constraints that do or don't hinder TEFA implementation.

Nevertheless, some themes are common. One of the most prevalent is a perceived tension between TEFA's emphasis on taking time to build solid, well-structured, transferable core subject knowledge, and the demands of state standards and high-stakes exams. This is the classic "breadth vs. depth" dilemma. Another is that many teachers have accumulated a large collection of curriculum materials, pedagogical methods, and instructional activities over time, are reluctant to abandon them completely, and are not sure whether or how they can integrate them with TEFA. As a result, many beginners sporadically insert TEFA interludes into their normal practice, despite warnings that it loses most of its effectiveness if used infrequently and inconsistently.

These five skill areas tend to demand TEFA learners' attention at different times, following a common sequence, but they are not *resolved* separately and in sequence. The skill areas are quite intertwined, so that progressing in one often requires attending to others. For example, as teachers wrestle with the problem of coaxing richer discussion from their students, they come to realize that the nature of the question being discussed is absolutely crucial; some questions simply do not support rich discussion. Thus, progress in orchestrating discussion becomes linked with progress in question design. Similarly, growth in a teacher's ability to figure out how students think aids the creation of better-tuned, more effective questions.

Professional Development Model

Since mastering TEFA requires developing a broad spectrum of skills, most teachers need extensive professional development (PD) support in order to succeed. In order to study how teachers learn TEFA, our current project includes a major PD component.

Many studies indicate that traditional teacher PD activities, in which subject or pedagogical content knowledge is "delivered", have little effect on teachers' practice (Bransford, Brown & Cocking, 1999). High quality science teacher PD should have the following characteristics: it should place teachers in the role of learners and "immerse [them] in inquiry, questioning, and experimentation"; it should be "both intensive and sustained"; it should "engage teachers in concrete teaching tasks and be based on teachers' experiences with students"; and it should "show teachers how to connect their work to specific standards for student performance" (Supovitz & Turner, 2000). Effective PD helps teachers learn continuously from practice through critical reflection and peer co-learning; targets classroom practices and student learning; focuses on the critical work of actual teaching rather than abstractions; forms a sustained scaffolding rather than a brief intervention; employs realistic case study and criticism methods; is integrated with daily teaching; ties pedagogy to content knowledge as well as to knowledge about student learning and cognition; and provides opportunities for meaningful teacher leadership (Ball & Cohen, 1999; NEIRTEC, 2004).

The project's PD program follows these known best practices for general science teacher PD. The program is sustained, lasting three years, and intensive. It begins with a four-day summer workshop designed to give teachers an overview of TEFA, solid experience with what TEFA "looks like", sufficient competence with the technology to begin using it, and some initial practice designing questions and conducting the question cycle. It continues with weekly after-school meetings for the first academic year. These focus on modeling TEFA for the participants, using a range of disciplines for sample content; developing theoretical frameworks to help participants with the TEFA skill areas; helping participants develop and critique CRS questions to use in their teaching; and sharing and discussing participants' TEFA-related experiences, concerns, and successes.

For the second and third academic years, PD switches to a new mode: meeting after school every three to four weeks, teachers engage in a type of collaborative action research called *enhanced normal practice* (Feldman, 1996; Feldman & Capobianco, 2000, Feldman & Minstrell, 2000). This consists of sharing and discussing anecdotes, trying out ideas, and pursuing systematic classroom inquiry (Feldman & Capobianco, in press), with the purpose of sustaining and extending teacher change and developing habits of ongoing inquiry into one's own practice (Ball & Cohen, 1999).

To provide teachers with a supportive community of co-learners, ensure school administrative support, and provide a more consistent learning experience for participants' students, our PD model calls for working with all, or at least a majority, of science teachers from one school. All PD activities are conducted on-site at the school.

Research Methods

School Sites

Three high schools within 45 minutes' drive of the University of Massachusetts Amherst have agreed to participate: that is, to have their entire science faculty take part in the TEFA PD program and to provide the necessary administrative support. The first school is in a rural high-needs district, with a largely homogeneous non-minority population. Science and math teachers from the high school and middle school, which are housed in the same building, are participating in the TEFA project. The second school is in a small residential city, is not high-needs, and serves a diverse population. The third school is one of the largest in Western Massachusetts, serving a broad suburban area on the edge of a city. A total of approximately 32 teachers are participating in the project.

Research Design

The project follows a mixed-methods, delayed-intervention, longitudinal repeated-measurements design. Two parallel strands of data acquisition are being pursued. One is the collection of field notes and records from PD activities for qualitative analysis. The second consists of the repeated application of several qualitative and quantitative instruments, beginning the year before the teachers' participation in TEFA begins, and continuing through the three years of PD. The three cohorts of teachers (three schools) are staggered, with PD beginning in August 2006, August 2007, and August 2008.

Data is collected with interviews, observations, and survey instruments of teachers and students. Several times each year every teacher's classes are videotaped. Audio recordings of PD course sessions and action research meetings are made and transcribed. Teachers are interviewed about their backgrounds and perspectives, their hurdles and concerns, and their lesson planning activities. In addition, students are surveyed twice a year about their perceptions of their classroom environment.

Data Analysis

Qualitative data are analyzed using methods described by Miles and Huberman (1994). Where applicable, frequency of practice and approximate time spent is tabulated. The data also illuminates other necessary skills that we had not previously identified. These are, essentially, "snapshots" of the teachers' practice: a partial profile on the skills mastered, those practiced but not mastered, and those not practiced. Analysis will assemble all of these profiles for a single teacher to construct a "trajectory" of the teacher's developing TEFA mastery. By comparing different teachers' trajectories, we hope to identify common patterns in teacher learning of TEFA. Comparison of data from treatment years with data from the baseline year will permit identification of immediate changes to teaching practice.

Preliminary Findings

As of January 2008, the TEFA project has completed 1.5 years of PD with the first cohort of teachers and one semester with the second. Our initial findings suggest that teachers' acceptance and use of the hardware, software, and pedagogic approach are dependent on a number of factors. These include:

1. Teachers' initial ability to use the CRS hardware and software. The ease with which the teachers were able to initially use the hardware and software had a significant effect on their learning curve for implementation of the TEFA approach.

2. Teachers' expectations of the abilities of their students. The teachers in the first year cohort included middle and high school math and science teachers. In general, the middle school teachers had much lower expectations of their students' ability to engage in quality discussions in response to TEFA items.

3. Teachers' initial feelings about approach. Although all the teachers in the first cohort were volunteers, some were more enthusiastic about the approach than others. Those teachers who began the project with less enthusiasm took up the approach much more slowly than those who were enthusiastic. In addition, teachers who began the project in agreement with the TEFA goals were more successful with the approach at an earlier time.

4. Comfort with "frontal teaching". The TEFA approach puts teachers in a position similar to frontal teaching. Although the goal is to have students actively engaged in the material, the class is orchestrated by the teacher. It appears that those teachers who can be characterized as excellent traditional teachers have more initial success with the approach. Those teachers who had less initial expertise with frontal teaching and had less ability with the use of hardware and software reported that they were rapidly left behind in the PD course.

5. Orientation towards reflection on practice. It appears that the teachers who were most successful with the approach during the first year were those who were self-reflective in the sense that when they questioned their practice, they focused on their beliefs and actions. Those

who were less successful tended to focus their reflection on how the students' abilities and attitudes affected their practice.

6. Beliefs about alignment of TEFA and the teachers' curriculum. In order for the TEFA approach to be successful, teachers need to use CRS with TEFA questions on a regular basis. Those teachers who believed that TEFA "fit" their curriculum used it more frequently; those who did not see a fit rarely used CRS and TEFA questions.

6. Summary

In this paper, we have described *Technology-Enhanced Formative Assessment* (TEFA), our pedagogical approach to teaching science with the assistance of a classroom response system. After summarizing the research base for our theoretical framework, we asserted that TEFA has been designed to help students develop expertise in science content, and to help prepare students for future learning. To do this, TEFA affirms four core principles, which we label *question-driven instruction*, *dialogical discourse*, *formative assessment*, and *meta-level communication*. These principles are enacted in the classroom via an iterative *question cycle*, aided by classroom response system technology.

We reported that teachers must develop skills in five different areas in order to master TEFA: operating the technology, designing questions, orchestrating classroom discourse, modeling students and adapting to their needs, and integrating TEFA with curricula and constraints. To help them do this, an intensive, sustained, collaborative, on-site professional development program is required. We described an ongoing project aimed at illuminating teachers' learning of TEFA and developing more effective TEFA professional development methods and curriculum.

For science teachers, implementing research-based best practices in science instruction — for example, creating a learning environment consistent with the four qualities identified in *How People Learn* (Bransford, Brown & Cocking, 1999) — can be difficult. Wholesale change, not isolated modifications or additions, is often called for, but this can be extremely intimidating and bewildering. We believe our pedagogical model does much to address this problem by integrating sound, comprehensive principles into a simple and approachable, yet flexible and rich, classroom methodology.

Acknowledgment

Preparation of this article and conduct of the research project described was/is supported by National Science Foundation (NSF) grant ESI-0456124. Development and investigation of the pedagogy described has been partially supported by NSF grants ESI-0456124, ESI-9730438, and DUE-9453881. Any opinions, findings, conclusions, and recommendations expressed herein are those of the authors and do not necessarily reflect the views of the NSF.

References

Abrahamson, L. A. (2006). A brief history of networked classrooms: Effects, cases, pedagogy, and implications. In Banks, D. A. (Ed.), *Audience Response Systems in Higher Education: Applications and Cases*. Idea Group Inc., Hershey, PA.

Ball, D. and Cohen, D. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In Sykes, G. and Darling-Hammond, L., editors, *Teaching as the Learning Profession: Handbook of Policy and Practice*, pages 3–32. Jossey Bass, San Francisco.

Banks, D. A. (Ed.) (2006). *Audience Response Systems in Higher Education: Applications and Cases*. Idea Group Inc., Hershey, PA.

Beatty, I. D. (2004). Transforming student learning with classroom communication systems. Research Bulletin ERB0403, Educause Center for Applied Research.

Beatty, I. D., Gerace, W. J., Leonard, W. J., and Dufresne, R. J. (2006a). Designing effective questions for classroom response system teaching. *American Journal of Physics*, 74(1):31–39.

Beatty, I. D., Leonard, W. J., Gerace, W. J., and Dufresne, R. J. (2006b). Question driven instruction: Teaching science (well) with an audience response system. In Banks, D. A., (Ed.), *Audience Response Systems in Higher Education: Applications and Cases.* Idea Group Inc., Hershey, PA.

Bell, B. and Cowie, B. (2001). The characteristics of formative assessment in science education. *Science Education*, 85(5):536–553.

Black, P., Harrison, C., Lee, C., Marshall, B., and Wiliam, D. (2002). *Working inside the black box: Assessment for learning in the classroom*. King’s College London Department of Education and Professional Studies, London, UK.

Black, P. and Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1):7–74.

Black, P. and Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2):139–148.

Black, P. and Wiliam, D. (2005). Developing a theory of formative assessment. In Gardner, J., (Ed.), *Assessment and Learning*. Sage Books.

Blum, L. (1999). Ethnicity, identity, and community. In Katz, M., Noddings, N., and Strike, K., editors, *Justice and Caring: The Search for Common Ground in Education*. Teachers College Press.

Bonwell, C. C. and Eison, J. A. (1991). Active learning: Creating excitement in the classroom. ASHE-ERIC Higher Education Report No. 1, ERIC Clearinghouse on Higher Education, The George Washington University.

Bransford, J. D., Brown, A. L., and Cocking, R. R. (1999). *How People Learn: Brain, Mind, Experience, and School*. National Academy Press, Washington, D.C.

Bransford, J. D. and Schwartz, D. (1999). Rethinking transfer: A simple proposal with multiple implications. In Iran-Nejad, A. and Pearson, P. D. (Eds.), *Review of Research in Education*, volume 24, pages 61–100. American Educational Research Association, Washington, DC.

Burnstein, R. A. and Lederman, L. M. (2003). Comparison of different commercial wireless keypad systems. *The Physics Teacher*, 41:272–275.

Carlsen, W. S. (2007). Language and science learning. In Abell, S. K. and Lederman, N. G. (Eds.), *Handbook of Research on Science Education*, chapter 3, pages 57–74. Lawrence Erlbaum Associates, Mahwah, NJ.

Chi, M. T. H. and Glaser, R. (1981). The measurement of expertise: Analysis of the development of knowledge and skills as a basis for assessing achievement. In Baker, E. L. and Quellmalz, E. S. (Eds.), *Design, Analysis and Policy in Testing*, pages 37–47. Sage Publications, Beverly Hills, CA.

Chin, C. and Chia, L.-G. (2004). Problem-based learning: Using students' questions to drive knowledge construction. *Science Education*, 88(5):707–727.

Davis, K., Feldman, A. P., Irwin, C., Pedevillano, E., Capobianco, B., and Weiss, T. (2003). Wearing the letter jacket: Legitimate participation in a collaborative science, mathematics, engineering, and technology education reform project. *School Science and Mathematics*, 103(3):121–134.

Dufresne, R. J., Gerace, W. J., Leonard, W. J., Mestre, J. P., and Wenk, L. (1996). Classtalk: A classroom communication system for active learning. *Journal of Computing in Higher Education*, 7:3–47.

Dufresne, R. J., Gerace, W. J., Mestre, J. P., and Leonard, W. J. (2000). ASK-IT/A2L: Assessing student knowledge with instructional technology. Technical Report Dufrense-2000aia, University of Massachusetts Scientific Reasoning Research Institute.

Dufresne, R. J., Leonard, W. J., and Gerace, W. J. (1993). A constructivist view of science education. Technical Report Dufresne-1993cvs, University of Massachusetts Scientific Reasoning Research Institute.

Dufresne, R. J., Thaden-Koch, T., Gerace, W. J., and Leonard, W. J. (2005). Knowledge representation and coordination in the transfer process. In Mestre, J. P., (Ed.), *Transfer of Learning from a Modern Multidisciplinary Perspective*, chapter 5, pages 89–119. Information Age Publishing.

Duncan, D. (2005). *Clickers in the classroom: How to enhance science teaching using classroom response systems*. Addison-Wesley, New York.

Elby, A. (2001). Helping physics students learn about learning. *American Journal of Physics Supplement*, 69(S1):S54–S64.

Fagen, A. P., Crouch, C. H., and Mazur, E. (2002). Peer instruction: Results from a range of classrooms. *The Physics Teacher*, 20:206–209.

Feldman, A. (1996). Enhancing the practice of physics teachers: Mechanisms for the generation and sharing of knowledge and understanding in collaborative action research. *Journal of Research in Science Teaching*, 33(5):513–540.

Feldman, A. P. and Capobianco, B. M. (2000). Action research in science education. Eric digest, ERIC Clearinghouse for Science, Mathematics, and Environmental Education, Columbus, OH.

Feldman, A. P. and Capobianco, B. M. (in press). Teacher learning of technology enhanced formative assessment. *Journal of Science Education and Technology.*

Feldman, A. and Minstrell, J. (2000). Action research as a research methodology for the study of the teaching and learning of science. In Kelly, E. and Lesh, R. (Eds.), *Handbook of Research*

Design in Mathematics and Science Education, pages 429–455. Lawrence Erlbaum Associates, Mahwah, NJ.

Fies, C. and Marshall, J. (2006). Classroom response systems: A review of the literature. *Journal of Science Education and Technology*, 15(1):101–109.

Gallagher, J. J. (2000). Teaching for understanding and application of science knowledge. *School Science and Mathematics*, 100(6):310–318.

Gerace, W. J. (1992). Contributions from cognitive research to mathematics and science education. In Grayson, D., (Ed.), *Workshop on Research in Science and Mathematics Education*, pages 25–44, Pietermaritzburg, South Africa. University of Natal.

Glaser, R. (1992). Expert knowledge and processes of thinking. In Halpern, D., editor, *Enhancing Thinking Skills in the Sciences and Mathematics*, pages 63–75. Lawrence Erlbaum Associates, Hillsdale, NJ.

Hammer, D. and Elby, A. (2003). Tapping epistemological resources for learning physics. *Journal of Learning Sciences*, 12:53–90.

Hammer, D., Elby, A., Scherr, R. E., and Redish, E. F. (2004). Resources, framing, and transfer. In Mestre, J. P. (Ed.), *Transfer of Learning: Research and Perspective.* Information Age Publishing, Greenwich, CT.

Hestenes, D., Wells, M., and Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, 30:141–158.

Koballa, T. R. and Glynn, S. M. (2007). Attitudinal and motivational constructs in science learning. In Abell, S. K. and Lederman, N. G. (Eds.), *Handbook of Research on Science Education*, chapter 4, pages 75–102. Lawrence Erlbaum Associates, Mahwah, NJ.

Larkin, J. (1979). Information processing models in science instruction. In Lochhead, J. and Clement, J. (Eds.), *Cognitive Process Instruction*, pages 109–118. Lawrence Erlbaum Associates, Hillsdale, NJ.

Lemke, J. L. (1990). *Talking Science: Language, Learning, and Values.* Ablex Publishing, Stamford, CT.

Leonard, W. J., Dufresne, R. J., Gerace, W. J., and Mestre, J. P. (1999). *Teacher's Guide to Accompany Minds-On Physics: Motion*, chapter B: Concept-based problem solving: Combining educational research results and practical experience to create a framework for learning physics and to derive effective classroom practices, pages B1–B26. Kendall/Hunt Publishing, Dubuque, IA.

Mazur, E. (1997). *Peer Instruction: A User's Manual*. Prentice Hall, Upper Saddle River, NJ.

McNeal, A. and D'Avanzo, C. (1997). *Student-Active Science: Models of innovation in undergraduate science teaching*. Saunders College Publishing, Orlando, FL.

Miles, M. and Huberman, M. (1994). *Qualitative Data Analysis*. Sage Publications, Thousand Oaks, CA, 2nd edition.

Mortimer, E. F. and Scott, P. H. (2003). *Meaning Making in Secondary Science Classrooms*. Open University Press.

NEIRTEC (2004). Professional development. Technical Brief 5. <http://neirtec.org/products/techbriefs/5.htm>.

Ohio State University Physics Education Research Group (2007). In-class polling methodology for effective education: The voting machines. <http://www.votingmachine.net/>.

Posner, G. J., Strike, K. A., Hewson, P. W., and Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66:211–227.

Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, 28(1):4–13.

Reay, N. W., Bao, L., Warnakulasooriya, R., and Baugh, G. (2006). Toward the effective use of voting machines in physics lectures. *American Journal of Physics*, 73(6):554–558.

Redish, E. F. (2003). A theoretical framework for physics education research: Modeling student thinking. In Vicentinni, M. and Redish, E. F. (Eds.), *Proceedings of the Varenna Summer School, "Enrico Fermi" Course CLVI*, Amsterdam. IOS Press.

Roschelle, J., Abrahamson, L. A., and Penuel, W. R. (2004a). Integrating classroom network technology and learning theory to improve classroom science learning: A literature synthesis. In *Annual Meeting of the American Educational Research Association*, San Diego, CA.

Roschelle, J. and Pea, R. (2002). A walk on the WILD side: How wireless handhelds may change computer-supported collaborative learning. *International Journal of Cognition and Technology*, 1:145–168.

Roschelle, J., Penuel, W. R., and Abrahamson, L. A. (2004b). Classroom response and communication systems: Research review and theory. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.

Sadler, R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18:119–144.

Scherr, R. E. (2007). Modeling student thinking: An example from special relativity. *American Journal of Physics*, 75(3):272–280.

Schwartz, D. L. and Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition & Instruction*, 22(2):129–184.

Scott, P., Asoki, H., and Leach, J. (2007). Student conceptions and conceptual learning in science. In Abell, S. K. and Lederman, N. G. (Eds.), *Handbook of Research on Science Education*, chapter 2, pages 31–56. Lawrence Erlbaum Associates, Mahwah, NJ.

Scott, P. H. and Mortimer, E. F. (2006). The tension between authoritative and dialogic discourse: A fundamental characteristic of meaning making interactions in high school science lessons. *Science Education*, 90(4):605–631.

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment FOR learning. *Phi Delta Kappan*, 83(10):758–765.

Supovitz, J. A. and Turner, H. M. (2000). The effects of professional development on science teaching practices and classroom culture. *Journal of Research in Science Teaching*, 37(9):963–980.

Sutton, C. (1998). New perspectives on language in science. In Fraser, B. J. and Tobin, K. G. (Eds.), *International Handbook of Science Education*, volume 1, pages 27–38. Kluwer Academic, Dordrecht, the Netherlands.

van Driel, J. H., Verloop, N., and de Vos, W. (1998). Developing science teachers' pedagogical content knowledge. *Journal of Research in Science Teaching*, 35(6):673–695.

Vygotsky, L. S. (1978). *The development of higher psychological processes*. Harvard University Press.

Vygotsky, L. S. (1987). Thinking and speech. In Rieber, R. W. and Carton, A. S. (Eds.), *The Collected Works of L. S. Vygotsky*. Plenum Press.

Wenk, L., Dufresne, R. J., Gerace, W. J., Leonard, W. J., and Mestre, J. P. (1997). Technology-assisted active learning in large lectures. In McNeal, A. and D'Avanzo, C. (Eds.), *Student-Active Science: Models of innovation in undergraduate science teaching*, pages 431–451. Saunders College Publishing, Orlando, FL.

Wertsch, J. V. (1991). *Voices of the Mind: A Sociocultural Approach to Mediated Action.* Harvard University Press.

Wiliam, D. (2007). Keeping learning on track: Classroom assessment and the regulation of learning. In Lester, F. K., editor, *Second Handbook of Mathematics Teaching and Learning*, pages 1051–1098. Information Age Publishing, Greenwich, CT.

Wilson, T. D. (2006). The power of social psychological interventions. *Science*, 313:1251–1252.

Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher education, 45*, 477-501.